



Apprentissage Statistique

Détection de structures communautaires dans des réseaux

Rédigé par

PRALON Nicolas

CÔME Olivier

SENE Assane

IMAG

INSTITUT MONTPELLIERAIN
ALEXANDER GROTHENDIECK

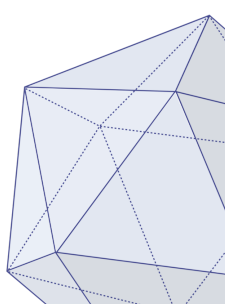


Table des matières

Introduction	2
Concept de Modularité	2

Introduction

De multiple réseaux, y compris les réseaux sociaux, les réseaux informatiques, se divisent plus ou moins naturellement en communautés. La détection de cette structure sous-jacente aux réseaux constitue un problème actuel, et de nombreuses approches ont été développées pour y répondre.

Dans ce rapport nous allons présenter une approche communément utilisée en apprentissage non supervisé, permettant de quantifier de la validité d'un partitionnement du réseau, les défaillances à cette approche et la mise en pratique des méthodes utilisées pour y répondre.

Concept de Modularité

L'étude d'éventuelle structures communautaires dans des réseaux peut formellement être présentée par l'étude de graphe. Ainsi nous considérons un réseau comme un graphe, et émettons certaines hypothèses à notre étude :

Soit $G = (V, E)$, un graphe tels que

$V = \{v_1, \dots, v_p\}$ l'ensemble des noeuds

$E \subset \{(v_i, v_j)_{i,j \in \{1, \dots, p\}} | i \neq j\} = V \times V$ l'ensemble des arêtes du graphe

G est une graph simple, non orienté, non pondéré, non labélisé.

Avant de décrire l'idée mise en oeuvre pour la détection de communautés, donnons quelques définitions.

Définition 1 (Densité). On appelle densité d'un graphe la valeur

$$D_G = \frac{|E|}{\frac{p^2 - p}{2}}$$

La densité d'un graph correspond à la fréquence d'arête dans le graph, il rend compte de la connexion entre les noeuds.

Définition 2 (Degré). On appelle degré d'un noeud i la valeur

$$d_i = |\{(v_i, v_j) \in E | j \in \{1, \dots, p\}\}|$$

et correspond au nombre de voisin du noeud i .

Définition 3 (Model nul). On appelle model nul d'un graph G , le graph G^* dont les $|E| = m$ arêtes ont été distribuées aléatoirement entre les noeuds de G